Cluster sampling to estimate breeding blackbird populations in North Dakota

Curtis O. Nelms, David L. Otis, George M. Linz, and William J. Bleier

Abstract Local blackbird populations cause most of the damage to sunflower crops in North Dakota. Precise estimates of blackbird breeding populations are necessary before management practices designed to reduce their numbers are implemented. We estimated populations of red-winged blackbirds (*Agelaius phoeniceus*), yellow-headed blackbirds (*Xanthocephalus xanthocephalus*), and common grackles (*Quiscalus quiscula*) in North Dakota from 798 quarter sections surveyed during May and June 1991. We tested the hypothesis that a stratified, 2-stage, cluster sampling scheme using Neyman allocation of cluster sample units would improve precision of statewide population estimates and be more cost-efficient than a stratified random sample design. Population estimates of breeding blackbirds were 1,425,000 (SE=43,000) pairs of red-winged blackbirds, 665,000 (SE=52,000) pairs of yellow-headed blackbirds, and 698,000 (SE=23,000) pairs of common grackles. Stratified cluster sampling using Neyman allocation yielded smaller variances than a stratified random sampling design would have with the same effort. The design effect ratio of 2.06 for all blackbirds indicates that effective sample size was about 400 random quarter sections as opposed to the 798 quarter sections actually surveyed. Because cluster sampling reduced overall cost/quarter section surveyed by 60% compared to stratified random sampling without reducing precision, cluster sampling was more cost-effective.

Key words cluster sampling, common grackle, North Dakota, population estimate, red-winged blackbird, stratification, yellow-headed blackbird

Sunflower is a major crop in North Dakota, with 793,212 hectares planted in 1998 (North Dakota Agricultural Statistics 1999). Hothem et al. (1988) estimated blackbird damage at $6.5 million in North Dakota in 1980. Red-winged blackbirds (*Agelaius phoeniceus*), yellow-headed blackbirds (*Xanthocephalus xanthocephalus*), and common grackles (*Quiscalus quiscula*), which compose about 9% of the avifauna in North Dakota (Igl and Johnson 1997), are primarily responsible for this damage (Linz and Hanzel 1997). Information on distribution and abundance of blackbirds is important for designing cost-effective integrated management programs to control sunflower losses to blackbirds (Leitch et al. 1997).

We estimated breeding blackbird populations in North Dakota in 1990 using 129 quarter sections (Nelms et al. 1994). We stratified the state into 4 strata, which represented its basic biotic regions. We allocated sample units to each stratum in pro-

Address for Curtis O. Nelms (deceased) and William J. Bleier: Department of Zoology, Stevens Hall, North Dakota State University, Fargo, ND 58105, USA. Address for David L. Otis: United States Department of Agriculture, Denver Wildlife Research Center, Building 16, Denver Federal Center, Denver, CO 80225, USA. Address for George M. Linz: United States Department of Agriculture, North Dakota Field Station, North Dakota State University, Fargo, ND 58105, USA. Present address for David L. Otis: South Carolina Cooperative Fish and Wildlife Research Unit, Clemson University, Clemson, SC 29634, USA. Present address for George M. Linz: Department of Agriculture, National Wildlife Research Center, Great Plains Field Station, 2110 Miriam Circle, Bismarck, ND 58102, USA.

portion to its area. Sample units, consisting of 1 legal quarter section, were then selected randomly and independently for each stratum (Figure 1). The population estimates made using these 129 quarter sections had large variances associated with them because the populations were clumped. Therefore, a refinement in sampling methodology might improve precision of the estimates. Large variances, probably from presence of clumped distributions, suggested that number of sampling units should be increased. Surveying more sampling units is desirable but more expensive, so we considered sampling designs other than stratified random sampling to reduce costs per unit sampled while retaining the statistical advantages of larger sample sizes.

Cost analysis, using time as the measure of effort invested, showed that approximately 4 quarter sections were surveyed for every team-day in 1990. Time required to survey quarter sections varied (range=110–156 min) among strata (Table 4). The average time spent on a quarter section in 1990 was 43 minutes; thus, only about 3 hours/day were spent surveying. Most of the survey time ($\bar{x}$=97 min/quarter section) was used traveling between and finding quarter sections. Therefore, a sampling design that reduced travel within the sampling periods could substantially increase the number of quarter sections that could be surveyed/team-day.

A 2-stage cluster sampling design, with townships as the primary units and quarter sections as secondary units, was selected for the 1991 survey. These sample units were found easily, which reduced time needed to organize the study and optimized the limited time available for surveying. Sufficient resources were available to survey a subsample of 10 quarter sections in each of 80 townships. Further, to improve efficiency of the survey, we used Neyman allocation (Cochran 1977:99), which minimizes the variance of a stratified random sample for a given cost (80 team-days), to allocate random samples to the 4 physiographic strata. Neyman allocation can be used when costs of surveying quarter sections (or any other sample unit) are fairly uniform across strata, total cost is fixed, and estimated variances within strata are different. Sample sizes necessarily increase to maintain precision as variability increases. The Neyman allocation method allows observers to distribute greater amounts of effort to less homogeneous strata, thus yielding smaller variance estimates for the population.

We used data collected in 1991 to test the hypothesis that a stratified, 2-stage cluster sampling scheme using Neyman allocation of cluster sample units would improve the precision of the statewide population estimates and be more cost-efficient than a stratified random sample design.

Methods

Cluster sampling

The sample units for this study were legal townships (36 sections, 9,324 ha) and the legal quarter section (160 acres, 64.75 ha). We used 2-stage cluster sampling, with townships as the primary units and quarter sections as the secondary units. We allocated 80 primary sample units to the 4 physiographic strata using Neyman allocation (Cochran 1977:98–99) and variances estimated from the 1990 survey. Within each stratum, we randomly

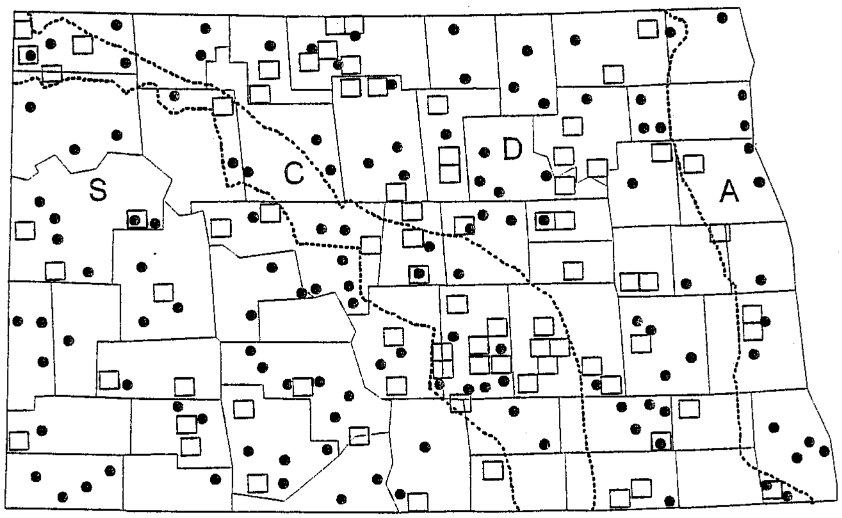


Figure 1. Biotic regions of North Dakota and distribution of random quarter-section sample units surveyed in 1990 (circles) and random township sample clusters (10 quarter sections from each cluster were sampled) surveyed in 1991 (boxes). Lettered strata with boundaries indicated by dashed lines designated as follows: (A) Agassiz Lake Plain, (B) Coteau Slope, (D) Drift Plain, (S) Slopes (after Besser 1985).

selected townships, with 4 in the Agassiz Lake stratum, 36 in the Drift Plain, 21 in the Missouri Coteau, and 19 in the Slopes (Figure 1). Within each township, we randomly selected 10 quarter sections. We surveyed quarter sections in the same manner as in 1990 (Nelms et al. 1994), except the quarter section was not divided into 4 transects and surveyed. Rather, clearly visible areas (e.g., plowed fields and grazed pastures) were thoroughly canvassed us-ing binoculars and spotting scopes, and enumerators walked or drove to and examined areas of reduced visibility, such as fence rows, potholes, brush and weeds, shelterbelts, ditches, and tall crop residues. We conducted all surveys during 2 daily sampling periods, from one-half hour after sunrise to 1100 hours and from 1600 hours to one-half hour before sunset. Two quarter sections in a township in southern McKenzie County were not surveyed because access was limited by steep slopes and poorly maintained roads.

Table 1. Population estimates and standard errors of red-winged blackbirds, yellow-headed blackbirds, common grackles, and the 3 species of blackbirds combined in North Dakota by species in 1990 and 1991, with *t*-tests comparing the 2 estimates.

Species	1990[a] Total	1990[a] SE	1991 Total	1991 SE	t	P
Red-winged blackbird	1,423,000	179,000	1,425,000	43,000	1.5	0.13
Yellow-headed blackbird	391,000	211,000	665,000	52,000	1.3	0.21
Common grackle	768,000	188,000	698,000	23,000	-0.4	0.71
Blackbirds	2,582,000	335,000	2,788,000	81,000	1.4	0.16

[a]Nelms et al. 1994

Analysis

We estimated population totals and variances using Cochran's (1977) statistical procedures for stratified random samples with optimum allocation and 2-stage cluster samples. A *t*-test for samples of unequal size compared the population estimates from 2-stage cluster sampling with Neyman allocation used in 1991 and the population estimates from random sampling with proportional allocation used in 1990 (Nelms et al. 1994). To better compare the designs and results of the 1990 and 1991 surveys, an adjusted estimate of the 1991 population variance was made. We estimated adjusted 1991 population variances by reducing the sample size values used in calculations to reflect number of team-days spent in 1990 while assuming within and among cluster variances

Table 2. Mean number (pairs) of red-winged blackbirds, yellow-headed blackbirds, common grackles, and the 3 species combined per quarter section within each stratum in North Dakota in 1991.

Stratum	$\bar{x}$	Standard deviation[a]: Within townships	Standard deviation[a]: Among townships	Cost associated with surveying[a]: Township	Cost associated with surveying[a]: Quarter section	Optimum number of quarters/ township[a]
Red-winged blackbird						
Lake Agassiz	4.0	6.1	5.5	126	30	5
Drift Plains	5.5	9.6	3.8	164	36	10
Missouri Coteau	8.6	13.2	11.7	133	33	7
Slopes	3.5	6.5	9.5	192	35	5
Yellow-headed blackbird						
Lake Agassiz	b	–	–	–	–	–
Drift Plains	4.8	20.1	15.5	164	36	9
Missouri Coteau	3.4	11.3	18.2	133	33	5
Slopes	0.1	0.7	0.03	192	35	10
Common grackle						
Lake Agassiz	1.3	2.0	0.1	126	30	11
Drift Plains	3.5	6.8	1.8	164	36	20
Missouri Coteau	3.1	5.6	0.4	133	33	7
Slopes	1.5	4.3	1.5	192	35	8
Blackbirds						
Lake Agassiz	5.2	6.6	7.0	126	30	5
Drift Plains	13.9	25.0	66.5	164	36	11
Missouri Coteau	15.1	21.3	19.0	133	33	5
Slopes	5.1	9.2	22.1	192	35	5

[a] Estimated from 1991 data
[b] No yellow-headed blackbirds were encountered on sample quarters in this stratum in 1991.

Table 3. The estimated population SE for 1991 with effort adjusted to be comparable to 1990 and percentage of change in SE compared to 1990 and 1991.

	Red-winged blackbird	Yellow-headed blackbird	Common grackle	Blackbirds
Adjusted 1991 SE	63,000	96,000	35,000	131,000
% change from 1990	-65	-55	-81	-61
% change from 1991	47	85	66	62

obtained in 1991 are good estimates of true variances (Cochran 1977). We calculated the design effect ratio and coefficient of intraclass correlations for each species in 1991 to compare the efficiency of stratified cluster sampling to stratified random sampling and estimate effective sample size (Kish 1965). The coefficient of intraclass correlation is the ratio of homogeneity of quarter sections within townships to the homogeneity of quarter sections in the state. The more homogeneous that quarter sections are within townships, then the greater the variance of the population estimate will be. The method of censusing a quarter used in 1991 reduced time actually spent on a quarter section by 20 minutes because strip transects were not used. This difference was subtracted from 1990 means to account for different survey techniques. We used the Wilcoxon rank sum test (Ott 1988) to compare costs of surveying samples using stratified cluster sampling in 1991 and stratified random sampling in 1990. Optimum number of quarter sections/township were estimated using 1991 data (Scheaffer et al. 1986), with time as a measure of cost.

Results

Survey

The 1991 survey results were similar to those obtained in 1990, with state-wide populations (pairs) of 1,425,000 red-winged blackbirds, 665,000 yellow-headed blackbirds, and 698,000 common grackles (Table 1). Density estimates of blackbirds, which are the basis of population estimates, varied with stratum (Table 2). The Missouri Coteau ($\bar{x}$=15.1/quarter section) and Drift Plains ($\bar{x}$=13.9/quarter section) had 2–3 times more blackbirds than did the Lake Agassiz ($\bar{x}$=5.2/quarter section) and Slopes ($\bar{x}$=5.1/quarter section) regions. Red-winged black-birds were more prevalent in the Missouri Coteau; densities of yellow-headed blackbirds and common grackles were greater in the Drift Plains.

To determine effects of the stratified cluster sampling with Neyman allocation design used in 1991, with greater effort (67 team-days), compared to the stratified random sampling with proportional allocation design used in 1990, we adjusted the 1991 estimated variances to reflect the reduced effort in 1990 (34 team-days). We estimated that if 34 team-days of effort had been used in 1991, standard errors would still have been reduced 55–81% from 1990 to 1991, which is attributable to using cluster sampling and Neyman allocation in 1991 (Table 3). There would have been a 47–85% increase in the standard error of the 1991 population estimates if 34 team-days had been expended rather than the 67 team-days actually used. This increase in variability is attributable to smaller sample sizes that would have resulted from a reduction in effort and illustrates the relationship between the level of sampling effort and variability in estimates.

Stratified cluster sampling yielded design effect ratios of 2.21, 1.85, 1.75, and 2.06 for red-winged blackbirds, yellow-headed blackbirds, common grackles, and blackbirds combined, respectively.

Table 4 . Mean time and standard error required to census quarter sections in 1990 and 1991, with adjusted means for 1990 to correct for a change in census technique and Wilcoxon rank sum tests comparing adjusted 1990 survey times and 1991 survey times.

	1990				1991				
Stratum	Mean (min.)	Adjusted mean[a]	SE	n	Mean (min.)	SE	n	z	P
Lake Agassiz	110	90	6	13	42	4	40	4.4	≤0.001
Drift Plains	130	110	11	50	49	2	210	8.2	≤0.001
Missouri Coteau	146	126	18	18	43	1	360	6.5	≤0.001
Slopes	156	136	11	48	53	3	188	9.1	≤0.001
North Dakota	140	120	6	129	48	1	798	14.8	≤0.001

[a] Quarters were surveyed an average of 20 minutes faster in 1991 because strip transects were not used; instead, observers went directly to areas of low visibility, such as shelterbelts, and canvassed high visibility areas, such as plowed fields, with binoculars and spotting scopes.

The coefficients of intraclass correlation ranged from 0.083 to 0.134.

Costs

Average cost of surveying a quarter section in North Dakota using cluster sampling and the modified complete census technique, measured in minutes of a team of 2 observers' working time, was 48 (SE=1) minutes. This cost includes a mean of 23 minutes spent censusing the quarter and an average of 25 minutes of travel time between townships. We prorated time spent traveling between 2 townships over the quarters sampled within the second township. Mean cost of surveying quarter sections in the 4 strata varied a small amount between strata (Table 4). Similar to 1990, time required to census a sample was greatest in the Missouri Slopes stratum (53 min, SE=3). The least time required/sample was in the Lake Agassiz Plain stratum (42 min, SE=4). The 1991 cluster sampling method resulted in a 60% reduction in the overall cost/quarter section, after we adjusted the 1990 data for the difference in census technique (z=14.8, 11,926 df, P<0.001). The reduction in cost/quarter was uniform across all strata.

Optimizing cluster samples

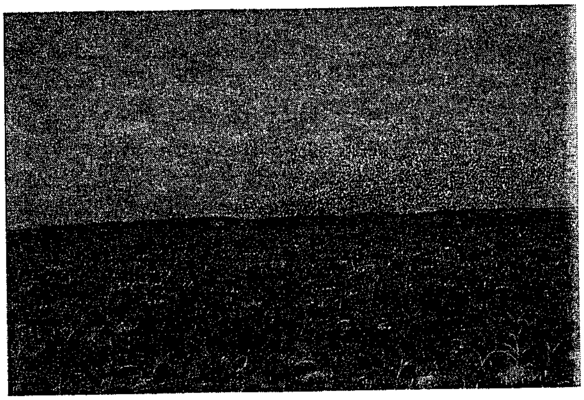
In 1991, we sampled 10 quarter sections (samples) in each township (cluster). Based on these data, the optimum number of quarter sections/township, which depends on the variances of quarter sections within townships and the variances among townships as well as survey costs and travel costs within and between townships, was greater than 10 for common grackles in the Drift Plain and the Lake Agassiz Plain strata and for the 3 species combined in the Drift Plain stratum (Table 2). Otherwise, the optimum number of quarter sections/township was 10 or less for each species. An optimum number of quarter sections needed to estimate the population of yellow-headed blackbirds in the Lake Agassiz Plain was not estimated because none were encountered on the quarter sections surveyed there. By using the optimum number of quarter sections in future surveys, a lower variance/unit cost should be achieved.

Discussion

The 1991 cluster sampling technique reduced the cost of doing individual quarter sections from 1990 levels, even after we adjusted for the modification of the survey method between years. Average cost/quarter section also was more uniform across strata in 1991 than in 1990. For example, cost of surveying a quarter section in the Missouri Slopes stratum was 46 minutes longer than in the Lake Agassiz stratum in 1990; the difference for the same areas in 1991 was only 11 minutes. Cost of surveying a quarter section in 1991 was 40% that of the 1990 adjusted cost, due primarily to reduced mean travel time between quarter sections.

The coefficients of intraclass correlation and design effects were used to estimate effect of cluster sampling versus random sampling without the confounding effects of Neyman allocation used in 1991 compared to the proportional allocation used in 1990 or the differences in effort between years. The coefficients of intraclass correlation indicate there was an effect of homogeneity among quarter sections within townships. The design effect ratios indicate that the effective sample size of the stratified cluster samples, compared to stratified random samples, was about 50% of the actual number. This means that we could expect comparable results whether sampling 400 quarter sections randomly or 80 clusters of 10 quarter sections each. A sample of 400 random quarter sections would lead to a small reduction in travel time spent between quarter sections as they would, on average, be closer together. On the other hand, time spent locating landowners and obtaining permission would not change, and a large proportion of time would still be expended during survey periods accessing random quarter sections, thus yielding costs very similar to those of 1990. Cluster sampling often enabled field researchers to use the midday period, from 1100 hours to 1600 hours, to seek access to all the sample quarter sections in a township and thus fully use survey periods to count birds. Because it takes 60% less time in the field to survey cluster samples, surveying twice as many quarter sections with the cluster sample method is more cost-effective and yields at least equivalent results.

Another way of looking at the relationship of cost differences between the 2 sampling methods is the number of quarter sections that can be surveyed per day's effort. Using stratified random sampling, a team of 2 observers surveyed 4 (but would have been 5 with the modified survey technique used in 1991) quarter sections/day in 1990 and the same team using the cluster sampling method averaged surveying 12 quarter sections/day in 1991. While



A flock of blackbirds ascends from a sunflower field.

the sampling effort in the field almost doubled from 1990 to 1991, number of quarter sections surveyed increased 6.2-fold. This doubling of effort would have meant about 335 quarter sections could have been surveyed using stratified random sampling (accounting for more efficient field survey technique used in 1991) when actually 798 quarter sections were surveyed using stratified cluster sampling with a resulting effective sample size of 400 stratified random quarter sections. A sample comparable to 400 stratified random quarter sections is expected to yield a more precise estimate of the population than 335 stratified random quarter sections surveyed for the same cost. A further consideration, whose cost was not tracked and thus not included in the calculations, was the logistics of preparing for and conducting field work (e.g., gathering maps and aerial photographs), which were simplified with cluster sampling.

Recommendations

For future blackbird population surveys in North Dakota, we recommend 1) 4 physiographic strata, 2) Neyman allocation of sample units, 3) cluster sampling with townships as the clusters and quarter sections as the samples, and 4) an optimum number of samples within each township should be sampled within each stratum, dependent on for which species the most precise estimates are needed (Table 2). In addition, blackbird populations should be surveyed periodically to facilitate clarification of the relationship between blackbird populations and sunflower damage.

The principles demonstrated in this paper of selecting between a cluster sample design or a random sample design based on cost-effectiveness can be applied to a variety of situations. In cases where cost of obtaining observations increases significantly as the distance separating sample sites increases, cluster sampling may be more economical. This is especially true when cluster sampling allows observers to maximize use of critical survey periods, such as with surveying interior forest breeding song birds (e.g., mornings) or intertidal zone invertebrates (e.g., during low tide). Examples of where cluster sampling may serve better than simple random sampling or stratified random sampling are the estimation of small mammal populations on alpine tundra, due to the great costs of accessing and surveying plots, and surveying plots for plant communities where access is time consuming due to fragmented land ownership or difficult terrain.

Acknowledgments. D. Bergman, D. Blixt, J. Homan, K. Scheidecker, B. Schank, M. Soehren, and S. Swanson collected field data. R. Carlson assisted with study design and statistical analyses. R. Christman and staff of the United States Department of Agriculture's Agricultural Stabilization and Conservation Service provided aerial photographs. J. Brammer, R. Carlson, J. Homan, and C. McMurl reviewed earlier drafts of this paper. Furthermore, we thank the private landowners and the managers of Theodore Roosevelt National Park (J. Bradybaugh), Upper Souris National Wildlife Refuge (D. Knauer), and J. Clark Salyer National Wildlife Refuge (R. Howard) for allowing us access to land. This research was supported by the United States Department of Agriculture, National Wildlife Research Center, Fort Collins, Colorado, and the Department of Zoology, North Dakota State University, Fargo, North Dakota.

Literature cited

Besser, J. F. 1985. Changes in breeding blackbird numbers in North Dakota from 1967 to 1981-82. Prairie Naturalist 17: 133-142.

Cochran, W. G. 1977. Sampling techniques. John Wiley and Sons, New York, New York.

Hothem, R. L., R. W. DeHaven, and S. D. Fairaizl. 1988. Bird damage to sunflower in North Dakota, South Dakota, and Minnesota, 1979–1981. United States Fish and Wildlife Service. Fish and Wildlife Technical Report 15.

Igl, L. D., and D. H. Johnson. 1997. Changes in breeding bird populations in North Dakota: 1967 to 1992–93. Auk 114:74–92.

Leitch, J. A., G. M. Linz, and Baltezore, J. F. 1997. Economics of cattail (*Typha* spp.) control to reduce blackbird damage to sunflower. Agriculture Ecosystems and Environment 65:141–149.

Linz, G. M., and J. J. Hanzel. 1997. Birds and sunflower. Pages

381–394 *in* A. Schneiter, editor. Sunflower Technology and Production. Agronomy Monograph 35. Second edition. American Society of Agronomy, Madison, Wisconsin.

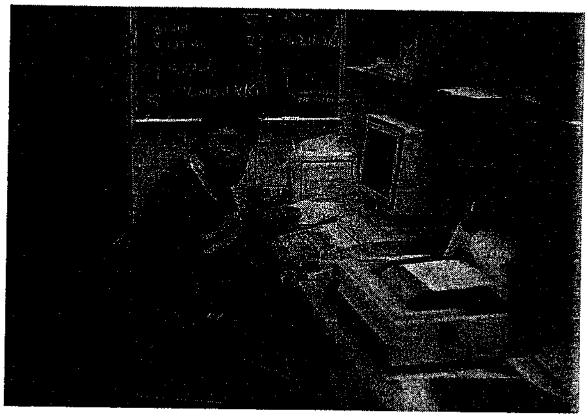
KISH, L. 1965. Survey sampling. John Wiley and Sons, New York, New York.

NELMS, C. O., W. J. BLEIER, D. J. OTIS, AND G. M. LINZ. 1994. Population estimates of breeding blackbirds in North Dakota, 1967, 1981–1982, and 1990. American Midland Naturalist. 132:256–263.

NORTH DAKOTA AGRICULTURAL STATISTICS SERVICE. 1999. North Dakota Agricultural Statistics 1999. North Dakota Agricultural Statistics Service, Fargo.

OTT, L. 1988. An introduction to statistical methods and data analysis. PWS-Kent, Boston, Massachusetts.

SCHEAFFER, R. L., W. MENDENHALL, AND L. OTT. 1986. Elementary survey sampling. Duxbury, Boston, Massachusetts.



Curtis O. Nelms (photo) passed away while enjoying a hunting trip in Utah. Curtis received his graduate education from North Dakota State University and was employed by the United States Fish and Wildlife Service at the time of his passing. We will miss Curtis's dry wit and unflappable personality. ***David L. Otis*** received his graduate education at Colorado State University. From 1977 to 1991, he was a research scientist and administrator at the Denver Wildlife Research Center. He has been unit leader of the South Carolina Cooperative Fish and Wildlife Research Unit since 1991. His research interests include animal population dynamics and statistical ecology. ***George M. Linz*** is a wildlife biologist and project leader with the National Wildlife Research Center at the Great Plains Field Station in Bismarck, North Dakota. He received B.S. and M.S. degrees in biology from Edinboro University of Pennsylvania and a Ph.D. in zoology from North Dakota State University. His current research focuses on the ecology of blackbirds in relation to ripening sunflower and on developing methods of reducing blackbird damage to sunflower. ***William J. Bleier*** received his graduate education at Texas Tech University. In 1975, he came to North Dakota State University, where he is now professor and chair of zoology. His current research interest is in integrated management of vertebrate pests, with particular emphasis on reducing avian damage to agricultural crops.

Associate editor: *Palmer*